

Identifying Nonconvexity in the Sets of Limited-Dimension Quantum Correlations

John Matthew Donohue^{1,*} and Elie Wolfe^{2,†}

¹*Institute for Quantum Computing and Department of Physics & Astronomy,
University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1*

²*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada, N2L 2Y5*

Quantum theory is known to be nonlocal in the sense that separated parties can perform measurements on a shared quantum state to obtain correlated probability distributions which cannot be achieved if the parties share only classical randomness. Here we find that the set of distributions compatible with sharing quantum states subject to some sufficiently restricted dimension is neither convex nor a superset of the classical distributions. We examine the relationship between quantum distributions associated with a dimensional constraint and classical distributions associated with limited shared randomness. We prove that quantum correlations are convex for certain finite dimension in certain Bell scenarios and that they sometimes offer a dimensional advantage in realizing local distributions. We also consider if there exist Bell scenarios where the set of quantum correlations is never convex with finite dimensionality.

I. INTRODUCTION

A poignant illustration of the non-classicality of quantum mechanics is found in Bell scenarios [1], where space-like-separated parties perform local measurements on a shared resource, as shown in Fig. 1. If the shared resource is an entangled quantum state, the conditional probability distribution resulting from the parties' measurements may be outside of the set of distributions achievable by sharing classical randomness or, equivalently, through a local hidden variable model (LHVM) [2, 3]. Per the device-independent formalism, we refer to any conditional probability distribution $p[abc\dots|xyz\dots]$ as a “box” [4–11], where the parties' choices of measurements $x, y, z\dots$ are referred to as “inputs” and their measurement results $a, b, c\dots$ as “outputs”. In this framework the quantum nature of the resource is assessed without information regarding the internal mechanism of the box, but rather by determining whether or not the box belongs to the LHVM-compatible set of boxes. Nonlocal boxes are those which cannot be implemented with classical shared randomness, and are highly valued as the resources which drive device-independent quantum communication and quantum cryptography [12–14].

For any given Bell scenario, the set of LHVM-compatible boxes is known as the local polytope; it has a finite number of facets given by Bell inequalities [2]. A box is defined as nonlocal if and only if it violates a Bell inequality. By contrast, we refer to the set of boxes implementable by sharing any quantum state as the quantum elliptope, which is also convex [15–18] but cannot be described in terms of linear inequalities [19, see 2, Fig. 4]. Boxes inside the quantum elliptope have conditional probabilities of the form

$$p[abc\dots|xyz\dots] = \text{Tr}[\rho (\hat{A}_{a|x} \otimes \hat{B}_{b|y} \otimes \hat{C}_{c|z} \dots)], \quad (1)$$

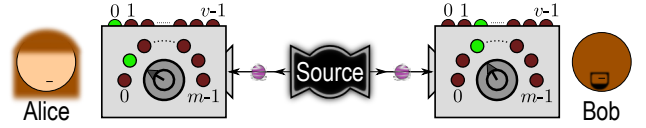


Figure 1. Bell scenarios may be viewed schematically as shown above, where n parties ($n = 2$ shown in this example) make synchronized measurements on systems prepared by a common source. Each party has m input (measurement setting) choices and v output (measurement result) possibilities. For $n = 2$, the parties are conventionally named Alice and Bob. Inputting x and y will return outputs a and b for Alice and Bob, respectively, with probability $p[ab|xy]$.

where the dimensions of ρ and the local measurement operators $\hat{A}_{a|x}, \hat{B}_{b|y}, \dots$ are unconstrained. To enforce no-signalling, i.e. to ensure that no information about the measurement choices of one party can be inferred by the measurement results of the other, the parties are assigned distinct Hilbert spaces [16, 20]. The operational boundary conditions delineating the quantum elliptope are notoriously hard to pin down, although various approximations to the quantum elliptope have recently become available [20–23]. The quantum elliptope strictly contains the local polytope; nevertheless boxes inside it cannot be used for nonlocal signalling [2].

In this letter, we show that convexity and containment of the local polytope are lost when the dimensions of the local Hilbert spaces are restricted. Although dimensionally constrained quantum correlations have already garnered considerable attention, most prior works have considered the parties as additionally having access to unlimited shared randomness. Examples include determining the degree to which Bell inequalities can be violated [24–26], assessing the security of a quantum cryptographic protocol [27–29], and testing if a given probability distribution is achievable [30–32]. The background assumption of unlimited shared randomness imposes convexity automatically, and obscures the fine-grained quantum-to-classical comparison which we seek. In contrast, our approach considers purely quantum systems or systems with limited shared randomness. Foundationally, the non-classical

* jdonohue@uwaterloo.ca

† ewolfe@perimeterinstitute.ca

nature of quantum mechanics is richer in detail if we compare quantum preparations to classical preparations at finite sizes, as opposed to only in the asymptotic limits of the quantum elliptope and local polytope [11]. The aim of this Letter is to explore this finer characterization.

Some important work has already been done in this framework. For example, Bowles *et al.* [33] recently re-analyzed prepare-and-measure signalling scenarios without the assumption of unlimited shared randomness [34]. The minimal classical or quantum dimension required to achieve some *unconditional* joint probability distribution, i.e. limiting each party to a single input, has recently been equated with the distribution's non-negative or positive semi-definite rank, respectively [35, 36], and an arbitrary strong quantum dimensional advantage has been noted [37]. For the usual multiple-inputs Bell scenario, the question of minimizing the quantum dimension required to implement a given box has recently been considered in Ref. [38], though see also Refs. [39, 40]. Pál and Vértesi [41] notably called attention to the nonconvexity of some dimensionally constrained quantum correlations. We resolve the open question they posed by showing that sharing merely qubits leads to a non-convex set of correlations in *any* non-trivial Bell scenario. We then extend the discussion to multipartite and multichotomous measurements. We prove that, in some scenarios, quantum correlations are guaranteed to be convex if the underlying Hilbert space dimension is sufficiently, but finitely, large.

II. NOTATION, DEFINITIONS, AND FUNDAMENTAL HIERARCHIES

We shall examine Bell scenarios where n space-like separated parties choose from among m inputs (measurement settings) each and, in response, observe among v possible outcomes (measurement results), as shown in Fig. 1. We label such Bell scenarios with the index $(n-m-v)$. The $(n-2-2)$ scenarios, for example, are those for which every party has access to two binary-outcome observables, and $(2-2-2)$ is the familiar CHSH nonlocality scenario [42, 43]. For analytical clarity, we consider exclusively symmetric Bell scenarios. As is conventional, for $(2-m-v)$ the two parties are referred to as Alice and Bob. We use x and y to indicate Alice and Bob's respective apparatus choices (box inputs), and use a and b to indicate their respective measurement outcomes (box outputs), starting all indices from 0.

To avoid over-specifying a no-signalling box, we parameterize boxes such that only the minimum number of conditional probabilities are specified to allow full reconstruction of the distribution. One such parameterization (inspired by Acín *et al.* [45]) is based on reserving-as-implicit all probabilities involving the outcome 0. For example, since $\sum_{b=0}^{v-1} p[ab|xy] = p[a|x]$ by no-signalling, one finds that $p[a0|xy] = p[a|x] - \sum_{b=1}^{v-1} p[ab|xy]$. In this parameterization scheme, each of the $\binom{n}{k} m^k$ possible k -partite input tuples one might condition upon has $(v-1)^k$

freely specifiable output probabilities. Consequently, any no-signalling box is specified in terms of precisely

$$\mathcal{F} = \sum_{k=1}^n \binom{n}{k} m^k (v-1)^k = (m(v-1) + 1)^n - 1 \quad (2)$$

total explicit parameters, regardless of the choice of parameterization scheme [2, 46].

For further simplification, we will only consider quantum measurement scenarios for which the local Hilbert spaces of every party have the same dimension¹, d . Let's denote the set of quantum boxes achievable with such a dimensional constraint as $\mathcal{Q}_d^{(n-m-v)}$. In this notation, therefore, the quantum elliptope is identically $\mathcal{Q}_\infty^{(n-m-v)}$. In this set, the correlation between the parties arises entirely from measurements on a shared quantum state, as shown in Fig. 2(a). We'll indicate $\mathcal{Q}_d^{(n-m-v)}$ via the shorthand \mathcal{Q}_d when the scenario $(n-m-v)$ is clear from context.

It will also be important for us to consider the set of boxes achievable by sharing only a classical random variable λ (as opposed to sharing a quantum state ρ), as seen in Fig. 2(b). In parallel with the notion of constrained local Hilbert space dimension, we let \mathcal{L} denote the set of boxes achievable by sharing only a classical random variable λ (as opposed to sharing a quantum state ρ). The set of classical boxes achieved using *constrained* shared randomness is analogously denoted $\mathcal{L}_{|\lambda|}$, where $|\lambda|$ refers to the dimension of the shared classical randomness, i.e. $\mathcal{L}_{|\lambda|}$ indicates that the parties share no more than $\log_2 |\lambda|$ classical bits in common. For example, if the parties share the outcomes of rolling two distinguishable and arbitrarily weighted dice, then $|\lambda| = 36$.

Boxes predicated on classical randomness are identically mixtures of product distributions², i.e. $\mathcal{L}_{|\lambda|}$ is the set of boxes for which

$$p[ab\dots|xy\dots] = \sum_{\lambda=0}^{|\lambda|-1} p[\lambda] p[a|x\lambda] p[b|y\lambda] \dots \quad (3)$$

Accordingly, \mathcal{L}_∞ may be thought of as the convex hull of \mathcal{L}_1 , where \mathcal{L}_1 is the set of boxes achievable without actually sharing any randomness. Indeed, that $\mathcal{L}_\infty = \text{ConvexHull}[\mathcal{L}_1]$ is a corollary of Fine's theorem [47–49].

While $\mathcal{L}_\infty = \text{ConvexHull}[\mathcal{L}_1]$ follows from Eq. (3), one can nevertheless span the local polytope without requiring infinite shared randomness. We define the minimum amount of shared randomness which allows every possible local box to be implemented as $|\lambda^*|$:

¹ The uniform-dimension constraint is insignificant in bipartite scenarios; see Appendix A for details.

² The notion of sharing classical randomness can be generalized to random variables with “memory”, but our analysis hews to the simplest model; see Ref. [2, Sec. II.G] for more details.

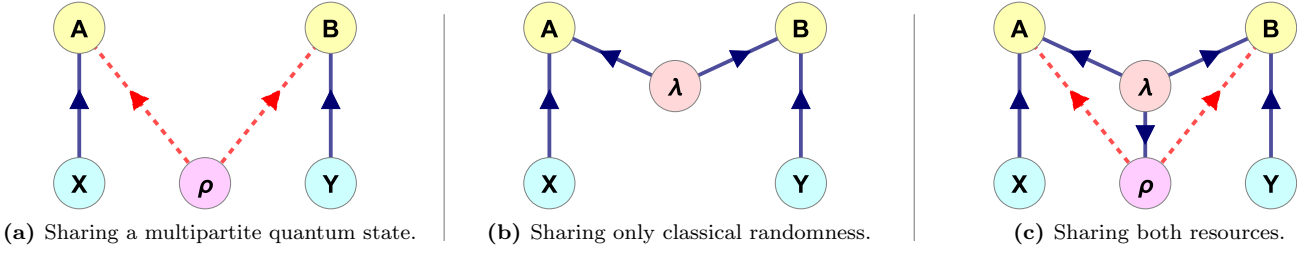


Figure 2. Various causal structures associated with two space-like separated parties who make synchronized measurements on systems prepared by a common source, as in Fig. 1. The joint distribution on the observed outcomes differs depending on the nature of the shared resource. If the preparation is quantum, such as in (a), then $p[ab|xy] = \text{Tr}[\rho \hat{A}_{a|x} \otimes \hat{B}_{b|y}]$, per Eq. (1). If the preparation is classical, such as in (b), then $p[ab|xy] = \sum_{\lambda} p[\lambda] p[a|x\lambda] p[b|y\lambda]$, per Eq. (3). If the shared system is both classical and quantum, such as in (c), then $p[ab|xy] = \sum_{\lambda} p[\lambda] \text{Tr}[\rho^{(\lambda)} \hat{A}_{a|x}^{(\lambda)} \otimes \hat{B}_{b|y}^{(\lambda)}]$, per Eq. (5). Hybrid preparations can equivalently be considered classical-quantum (cq) states [40, 44].

Definition $|\lambda^*|$: The classical Carathéodory number $|\lambda^*|$ is the minimum amount of classical randomness that must be shared in order to span the local polytope. Formally,

$$|\lambda^*| \equiv \min |\lambda| \text{ s.t. } \mathcal{L}_{\infty}^{(n-m-v)} = \mathcal{L}_{|\lambda|}^{(n-m-v)} \quad (4)$$

where $|\lambda^*|$ intrinsically depends implicitly on $(n-m-v)$. We call $|\lambda^*|$ the classical Carathéodory number [50, 51] of \mathcal{L}_1 because every $P \in \text{ConvexHull}[\mathcal{L}_1]$ can be decomposed as a convex mixture of at-most $|\lambda^*|$ boxes from \mathcal{L}_1 , whereas $|\lambda^*| - 1$ would not be adequate. Details regarding the determination of $|\lambda^*|$, including the result that $|\lambda^*| = 4$ for the (2-2-2) scenario, can be found in Appendix B [52].

Note that boxes arising from measurements on *separable* quantum states are equivalent to boxes arising from shared classical randomness and vice versa. Critically, the set of all classical boxes with degree of shared classical randomness $|\lambda|$ is contained by the set of boxes implemented quantumly by sharing separable qudits with local Hilbert space dimension $d \geq |\lambda|$. Formally,

Axiom (1): $\mathcal{L}_{|\lambda|} \subseteq (\mathcal{Q} : \text{sep})_{d \geq |\lambda|}$, i.e. any box that can be constructed by sharing a classical dit can also be composed by sharing qudits.

Corollary (1): $\mathcal{L}_{\infty} \subseteq \mathcal{Q}_d$ whenever $|\lambda^*| \leq d$.

Corollary (2): If $\mathcal{L}_{\infty} \not\subseteq \mathcal{Q}_d$, then \mathcal{Q}_d is not convex.

Proof. Every $P \in \mathcal{L}_{|\lambda|}$ can be mapped to a $P' \in (\mathcal{Q} : \text{sep})_{d=|\lambda|}$ by the following construction on Eq. (1): $\rho \rightarrow \sum_{\lambda}^{|\lambda|} p[\lambda] (|\lambda\rangle\langle\lambda|)^{\otimes n}$, $\hat{A}_{a|x} \rightarrow \sum_{\lambda}^{|\lambda|} p[a|x\lambda] |\lambda\rangle\langle\lambda|$, and so on. Thus, for example, any box in \mathcal{L}_1 can be implemented quantumly by sharing a product state. Note that the shared randomness required to enclose the separable quantum set of dimension d is $|\lambda| \geq d^n$, as shown in Appendix C [53, 54].

The corollaries follow since $\mathcal{L}_{\infty} = \mathcal{L}_{|\lambda^*|} = \text{ConvexHull}[\mathcal{L}_1]$. It follows from Eq. (4) and **Ax. (1)** that the local polytope is contained by the quantum set with Hilbert space dimension $d \geq |\lambda^*|$. Since $\mathcal{L}_{\infty} = \text{ConvexHull}[\mathcal{L}_1]$, we conclude that \mathcal{Q}_d is not convex if it does not contain the local polytope, i.e. whenever

$\mathcal{L}_{\infty} \not\subseteq \mathcal{Q}_d$. Therefore, evidence of any LHV boxes not achievable by sharing quantum states is evidence of non-convexity. \square

We conjecture that the shared quantum state must have local dimension $d \geq |\lambda^*|$ in order to contain the local polytope, i.e. the converse of **Cor. (1)**, that $\mathcal{L}_{\infty} \not\subseteq \mathcal{Q}_{d < |\lambda^*|}$. For that matter, we conjecture that the converse of **Ax. (1)** is also true, but this is a fundamental unproven open question.

Conjecture 1: $\mathcal{L}_{|\lambda| \leq |\lambda^*|} \not\subseteq \mathcal{Q}_{d < |\lambda|}$, i.e. the set of boxes which may be realized by sharing (possibly entangled) qudits of local Hilbert space dimension d is conjectured to *not* contain the set of boxes constructable by sharing a classical random variable of dimension $|\lambda| \leq |\lambda^*|$ if $d < |\lambda|$.

Note that **Conj. 1** would imply the nonconvexity of \mathcal{Q}_d whenever $d < |\lambda^*|$, per **Cor. (2)**. One can use entropic measures to infer a trivial lower bound on the smallest d for which $\mathcal{L}_{|\lambda|} \subseteq \mathcal{Q}_d$ by comparing the largest mutual information which can be mediated through the given $|\lambda|$ to the maximum total correlation capacity of quantum states of dimension d [56, Eq. (16)].

III. NON-CONVEX SETS OF QUANTUM CORRELATIONS

With the above definitions in place, we next demonstrate the nonconvexity of the set of quantum distributions explicitly. For this purpose, it suffices to focus on the most trivial example imaginable: a one-dimensional quantum system, ergo the set \mathcal{Q}_1 .

Proposition (1): \mathcal{Q}_1 is not convex.

Proof. If $d = 1$ then the only possible quantum state that Alice and Bob can “share” is the product state $|0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B$. No matter how Alice and Bob choose their POVM elements, their joint probability distributions will always factorize to a product distribution, $p[ab|xy] = \text{Tr}[\hat{A}_{a|x}] \text{Tr}[\hat{B}_{b|y}] = p[a|x] \times p[b|y]$. Indeed, \mathcal{Q}_1 is the set of all product distributions, equivalent to

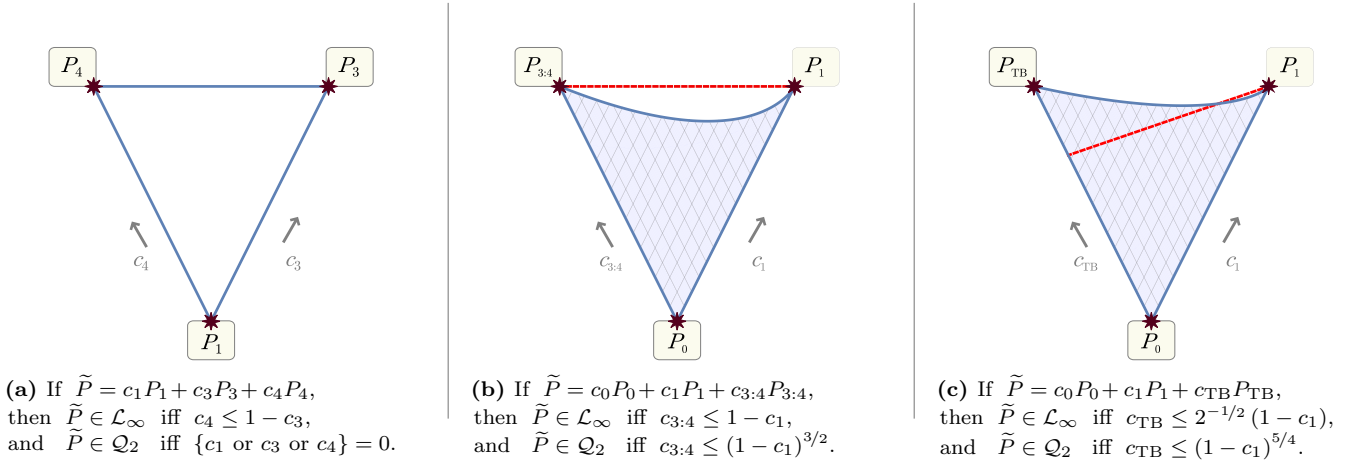


Figure 3. Three examples of the nonconvexity of the set \mathcal{Q}_2 are shown above. Each triangle represents the set of possible convex combination of its vertex boxes as defined in Table I, where c_i represents the weight of P_i in the mixture and $\sum_i c_i = 1$. To emphasize the notion of convex combination we have plotted the regions as equilateral triangles; nevertheless any pair of edges should be thought of as two independent axes such that the third weight is fixed by normalization. Although in each triangle all the vertex boxes are achievable with shared-qubits preparations, only a fraction of possible convex combinations are also achievable, as indicated by the shaded region. Indeed, in (a) only the zero-area edges of the triangle are qubit-achievable. The local polytope \mathcal{L}_∞ contains the entire triangular regions of (a) and (b) since the vertex boxes of those triangle are all local, and in (c) only the region below the dashed red line is consistent with classical shared randomness. The essential feature of all three triangles is that the shaded subset representing \mathcal{Q}_2 is not convex and that $\mathcal{L}_\infty \not\subseteq \mathcal{Q}_2$. The \mathcal{Q}_2 region in (c) was obtained via computational maximization; $(1 - c_1)^{5/4}$ was noticed to precisely coincide with the numeric boundary. The given boundaries of \mathcal{Q}_2 in (a) and (b) were recovered analytically. Consistent with Ref. [41], the qubit-accessible region shrinks if one restricts to projective measurements. The edges where $c_1 \neq 0$ in (a), for example, are not in \mathcal{Q}_2 -PVM, and in (c) we find that $\tilde{P} \in \mathcal{Q}_2$ -PVM only if $c_{TB} \lesssim (1 - c_1)^{3/2}$. Thus the use of POVMs can be certified in a device-independent way given a dimension promise, akin the the results of Ref. [55].

the set the of boxes achievable with any d subject to constraining ρ to be a product state. As such, ρ isn't really "shared" at all; the only randomness is local noise. These restrictions are certainly limiting, but because all local deterministic boxes are achievable in \mathcal{Q}_1 , we thus have $\mathcal{L}_\infty \subseteq \text{ConvexHull}[\mathcal{Q}_1]$.

	$\langle A_0 \rangle$	$\langle A_1 \rangle$	$\langle B_0 \rangle$	$\langle B_1 \rangle$	$\langle A_0 B_0 \rangle$	$\langle A_1 B_0 \rangle$	$\langle A_0 B_1 \rangle$	$\langle A_1 B_1 \rangle$
P_0	0	0	0	0	0	0	0	0
P_1	1	1	1	1	1	1	1	1
P_2	-1	-1	-1	-1	1	1	1	1
P_3	1	-1	1	-1	1	-1	-1	1
P_4	-1	1	-1	1	1	-1	-1	1
$P_{3:4}$	0	0	0	0	1	-1	-1	1
$P_{1:4}$	0	0	0	0	1	0	0	1
P_{TB}	0	0	0	0	$2^{-1/2}$	$2^{-1/2}$	$2^{-1/2}$	$-2^{-1/2}$

Table I. A list of bipartite conditional probability distributions, or boxes, in the (2-2-2) scenario. Per convention [2], we parameterize binary-output boxes in terms of outcome biases, i.e. $\langle A_x \rangle = p[a=1|x] - p[a=0|x]$ and $\langle A_x B_y \rangle = p[a=b|xy] - p[a \neq b|xy]$. Note that the five boxes P_0 through P_4 are product distributions, achievable even absent any shared randomness, i.e. $|\lambda| = 1$. $P_{i:j}$, by contrast, indicates the equally-weighted mixture of boxes P_i through P_j , requires non-trivial shared randomness. Only P_{TB} , the quantum box which achieves the Tsirelson bound [43, 57], is nonlocal. Every box in this table is achievable with qubits, but many mixtures of these boxes *cannot* be achieved using qubits, per Fig. 3.

Suppose Alice and Bob both have access to a uniformly-distributed variable $0 \leq \lambda \leq v - 1$, and let their marginal probabilistic dependencies on λ be such that $p[a|x\lambda] \rightarrow \delta[a = \lambda]$ and $p[b|y\lambda] \rightarrow \delta[b = \lambda]$ regardless of x or y . The resulting box $P \in \mathcal{L}_v$ is such that $p[\lambda\lambda|xy] = p[\lambda|x] = p[\lambda|y] = v^{-1}$ where P does not factorize, i.e. $p[\lambda\lambda|xy] \neq p[\lambda|x] \times p[\lambda|y]$. As such, $P \in \mathcal{L}_\infty$ yet $P \notin \mathcal{Q}_1$. By **Cor. (2)**, then, \mathcal{Q}_1 is not convex. \square

In a less trivial example, we also show that nonconvexity persists in the set \mathcal{Q}_2 when the number of inputs $m \geq 2$. In Ref. [41] the nonconvexity of \mathcal{Q}_2 was also proven, but only for scenarios with $m \geq 4$.

Proposition (2) : \mathcal{Q}_2 is not convex (for all $m \geq 2$).
For example, in the (2-2-2) scenario, the set of boxes that can be achieved by sharing qubits is nonconvex.

Proof. To show this, we considered maximally general two-qubit preparations and measurement schemes, and found that convex combinations of various boxes in \mathcal{Q}_2 listed in Table I were not themselves achievable with qubits³; see Fig. 3 for illustrative examples. We initially

³ A brute-force proof seemed necessary to the authors. The method of Sikora *et al.* [38], however, can also be used to prove the essential nonconvexity, by certifying that *some* of the boxes in the unshaded region of Fig. 3(a) necessitate a quantum description

identified a few nonconvex boundaries of \mathcal{Q}_2 via computational maximization, most of which we subsequently recovered analytically. Details of our state and measurements parametrization may be found in Appendix D [17, 55, 58–65]. The full derivation of the boundaries indicated in Fig. 3 may be found in a *Mathematica*™ notebook in the Supplementary Online materials [66]. \square

Proposition (3) : $(\mathcal{L}_\infty \setminus \mathcal{L}_{|\lambda|=d}) \cap \mathcal{Q}_d \neq \emptyset$, i.e. there exist local boxes which classically require shared randomness of dimension at least $|\lambda|$, but which admit quantum “shortcuts” through implementations using quantum systems of smaller dimension $d < |\lambda|$.

Proof. Consider the local box $P_{1:4}$ from Table I. If Alice and Bob choose the same input then their outputs are perfectly correlated, but if they choose different inputs then their outputs are unrelated. To achieve $P_{1:4}$ classically requires flipping two coins: When Alice or Bob input 0 their output is given by the first coin flip. The input 1, however, returns the second coin flip. Thus $P_{1:4}$ requires $|\lambda| \geq 4$, i.e. $P_{1:4} \in \mathcal{L}_\infty \setminus \mathcal{L}_3$. Alternatively, $P_{1:4}$ is quantumly achievable by sharing the entangled pure state $\frac{|00\rangle + |11\rangle}{\sqrt{2}}$ and letting $\hat{A}_{1|0} - \hat{A}_{0|0} = \hat{B}_{1|0} - \hat{B}_{0|0} = \hat{\sigma}_Z$ and $\hat{A}_{1|1} - \hat{A}_{0|1} = \hat{B}_{1|1} - \hat{B}_{0|1} = \hat{\sigma}_X$. So, $P_{1:4} \in \mathcal{Q}_2$. \square

Although a box which requires $|\lambda| \geq 4$ is inside \mathcal{Q}_2 , nevertheless $\mathcal{L}_4 \not\subseteq \mathcal{Q}_2$. Indeed, the empty interior of Fig. 3(a) shows that $\mathcal{L}_3 \not\subseteq \mathcal{Q}_2$, as is expected per **Conj. 1**. Thus entanglement enables not only non-locality, i.e. boxes forbidden classically, but also *super*-locality: quantum measurement schemes can occasionally reproduce boxes of higher corresponding shared randomness dimension, effectively simulating $|\lambda| > d$. See Zhang [37, Sec. 4.1] for a thorough discussion of both the existence and extent of super-locality in the special case (2-1- v). Interestingly, super-locality can occur even in the absence of entanglement [67].

IV. REGAINING CONVEXITY BY ADDING SHARED RANDOMNESS

The quantum elliptope, i.e. \mathcal{Q}_∞ , is convex. The convexity of the quantum elliptope, i.e. \mathcal{Q}_∞ , is well established; see for example Refs. [15, Sec. 5C] and [16, Sec. 5]; proof may also be given in terms of properties of C*-algebras [17, 18]. We have shown in **Prop. (2)**, however, that for finite local Hilbert space dimension \mathcal{Q}_d is sometimes not convex. The nonconvexity of \mathcal{Q}_d was also demonstrated by Pál and Vértesi [41]. There is no contradiction between the convexity of \mathcal{Q}_∞ and the nonconvexity of \mathcal{Q}_d ; rather,

convexification of quantum boxes requires either classical shared randomness or, equivalently, comes at the expense of increasing the local Hilbert space dimension.

The reason quantum boxes cannot be mixed without increasing Hilbert space dimension is because the measurements are local. The local nature of the measurement operators means that the composite POVM element associated with some global input and output $\hat{M}_{ab\dots|xy\dots} = \hat{A}_{a|x} \otimes \hat{B}_{b|y} \otimes \dots$ is necessarily a product operator. Mixtures of product operators, sometimes known as separable superoperators [68, 69], are generally no longer product operators.

On the other hand, access to classical randomness allows for the quantum preparations and measurements to co-depend on a shared classical hidden variable. Indeed, any combination of N qudit-based boxes can be implemented by sharing a single qudit, so long as the qudit is prepared according to a classical variable λ of dimension N , and the variable λ remains accessible to the measurements, i.e. per Fig. 2(c). Explicitly, we imagine the quantum preparations and measurements to co-depend deterministically on λ , i.e. $p[\rho^{(\gamma)}|\lambda] = p[\hat{A}_{a|x}^{(\gamma)}|\lambda] = \dots = \delta[\gamma - \lambda]$. By this construction, the resulting hybrid quantum-classical boxes yields

$$p[ab\dots|xy\dots] = \sum_{\lambda=0}^{|\lambda|-1} p[\lambda] \text{Tr}[\rho^{(\lambda)} \hat{A}_{a|x}^{(\lambda)} \otimes \hat{B}_{b|y}^{(\lambda)} \dots]. \quad (5)$$

Let’s denote the set of boxes achievable using quantum systems of dimension d with the assistance of shared randomness of dimension $|\lambda|$ as $\mathcal{Q}_d + \mathcal{L}_{|\lambda|}$. A hybrid box can be converted into a purely quantum description by embedding the shared randomness into a block-diagonal classical-quantum (cq) state of dimension $d \times |\lambda|$ [40, 44, 70, 71], and as such:

Axiom (2) : $\mathcal{L}_{|\lambda|} + \mathcal{Q}_d \subseteq \mathcal{Q}_{d' \geq d \times |\lambda|}$, i.e. any hybrid quantum-classic box can be thought of as a purely-quantum box predicated on a sufficiently larger local Hilbert space dimension $d' = d \times |\lambda|$.

Proof. To quote Baez [70]: “Roughly speaking, if we have a physical system whose states are either states of $\rho^{(i)}$ OR states of $\rho^{(j)}$, its Hilbert space will be the direct sum $\mathcal{H}^{(i)} \oplus \mathcal{H}^{(j)}$.” The direct sum of N Hilbert spaces each of dimension d can be embedded in a single Hilbert space of dimension $d \times N$ [71]. See Appendix E for further details. \square

Having shown that \mathcal{Q}_d can be non-convex and having established how to include finite shared randomness into our framework, we next ask how much supplementary shared randomness is required so that $\mathcal{L}_{|\lambda|} + \mathcal{Q}_d$ will be convex. Equivalently, what is the worst-case number of different $P \in \mathcal{Q}_d$ that must be combined in order to simulate *any* box which is expressible as a mixture of boxes from \mathcal{Q}_d ? We rephrase that question into a definition.

in terms of a Hilbert space dimension greater than two, i.e. that qubits would be incompatible. Brute-force feasibility-checking is nevertheless required to generate the nonconvexity illustrations in Fig. 3.

Definition $|\lambda_Q^*|_d$: The Quantum Carathéodory Number $|\lambda_Q^*|_d^{(n-m-v)}$ is the minimum amount of supplementary classical correlation that must be supplied for $Q_d^{(n-m-v)}$ to become completely convexified. Formally,

$$|\lambda_Q^*|_d \equiv \min |\lambda| \text{ such that } \mathcal{L}_{|\lambda|} + Q_d^{(n-m-v)} = \text{ConvexHull}[Q_d]. \quad (6)$$

We call $|\lambda_Q^*|_d$ the quantum Carathéodory number because Eq. (6) dictates that every $P \in \text{ConvexHull}[Q_d]$ can be decomposed as a convex mixture of at-most $|\lambda_Q^*|_d$ boxes from Q_d , whereas $|\lambda_Q^*|_d - 1$ would not be adequate.

Although $|\lambda_Q^*|_d$ may depend on d , there is a way to upper-bound the quantum Carathéodory number independently of d .

Theorem (1) : $|\lambda_Q^*|_d \leq (m(v-1) + 1)^n - 1$.

Proof. A 1929 theorem of Werner Fenchel [72, 73] states that any point within the convex hull of a not-necessarily-convex \mathcal{F} -dimensional closed and pathwise-connected set can be decomposed as a convex mixture of at-most \mathcal{F} points in the set; Fenchel's theorem is a strengthening of Carathéodory's theorem [51] for the special case of indivisible sets such as pathwise-connected ones. Q_d is a pathwise-connected closed compactum of dimension \mathcal{F} ; pathwise-connectedness follows from the continuous parameterization of both states and measurements and the Intermediate Value Theorem. The statistical space dimension of no-signalling boxes in a symmetric Bell scenarios is given by Eq. (2) [2, 46]. \square

For example, any box in $\text{ConvexHull}[Q_d^{(2-2-2)}]$ can be expressed as the convex combination of at-most eight qudit-based boxes. The actual quantum Carathéodory number may potentially be much lower than this upper bound, and is likely to have an explicit dependence on d . All we've established from **Prop. (2)** is that $|\lambda_Q^*|_2 \neq 1$. A tighter upper bound for $|\lambda_Q^*|_d$ than is given in **Thm. (1)** is desideratum for future research.

Importantly, the finiteness of $|\lambda_Q^*|_d$ can be used to guarantee the convexity of certain quantum sets $Q_d^{(n-m-v)}$ for finite dimension d .

Theorem (2) : If $Q_\infty = \text{ConvexHull}[Q_d]$ for a given Bell scenario, i.e. all extremal quantum distributions are achievable with qudits of dimension d , then $Q_{d'}$ is convex if $d' \geq d \times |\lambda_Q^*|_d$.

Corollary (3) : For $(n-2-2)$ scenarios, $Q_{d'}$ is convex if $d' \geq 2(3^n - 1)$, or assuming the upper bound of **Thm. (1)** is not tight, convexity is certain whenever $d' \geq 2|\lambda_Q^*|_2$.

Proof. By the definition of the quantum Carathéodory number, $\text{ConvexHull}[Q_d] = \mathcal{L}_{|\lambda_Q^*|_d} + Q_d$. By **Ax. (2)** we can embed the supplementary classical correlations into the shared quantum state such that $\text{ConvexHull}[Q_d] \subseteq Q_{d'}$ whenever $d' \geq d \times |\lambda_Q^*|_d$. However, the promise

of $Q_\infty = \text{ConvexHull}[Q_d]$ means that $Q_\infty \subseteq Q_{d'}$, from which it follows that $Q_\infty = Q_{d'}$. Thus, $Q_{d'}$ is guaranteed to be convex.

The corollary is a consequence of Masanes' theorem [74, 75], which states that (projective) measurements on merely shared qubits are capable of achieving all extremal quantum distributions for scenarios involving two binary measurements per party [76]. Masanes' theorem is often cited when noting that the maximum violation of Bell inequalities for such scenarios can be computed by maximizing over qubit-based boxes. \square

It is not clear if **Cor. (3)** can be extended to more general scenarios. For example, the I_{3322} Bell inequality in the $(2-m-2)$ many-measurement-choices scenario is apparently ever-more-violated as d is increased [77–80]. The boundaries of the quantum set increasing with dimension prevent using the above arguments, and indeed it remains an open question whether $Q_d^{(2-3-2)}$ is convex for *any* finite d . On the other hand, there is numerical evidence that $(2-2-v)$ many-measurement-outcomes scenarios, such as the CGLMP scenario [81], achieve maximum nonlocality at $d = v$ [22, 82]. The maximum violation of every Bell inequality, however, does not necessarily imply that all quantum extremal distributions have been achieved, so it is not certain that convexity is achieved. See Ref. [2, Sec. III.B] for further details.

Conjecture 2 : We conjecture that nonconvexity is nothing more than an artifact of not spanning the local polytope, i.e. that if $\mathcal{L}_{|\lambda^*|} \subseteq \mathcal{L}_{|\lambda|} + Q_d$ then $\mathcal{L}_{|\lambda|} + Q_d$ is convex.

In stating **Conj. 2** we used the equivalence $\mathcal{L}_\infty = \mathcal{L}_{|\lambda^*|}$ per Eq. (4). Recall that the converse of **Conj. 2** is obviously true: if a set of boxes – quantum, classical, or hybrid – does not contain the local polytope, then the set is not convex per **Cor. (2)**.

Conj. 2 amounts to speculating that $|\lambda_Q^*|_d = \min |\lambda|$ such that $\mathcal{L}_{|\lambda^*|} \subseteq \mathcal{L}_{|\lambda|} + Q_d$. If true, this would replace **Thm. (1)** with the claim $|\lambda_Q^*|_d \leq \lceil |\lambda^*|/d \rceil$. This follows from $\mathcal{L}_d \subseteq Q_d$ per **Ax. (1)** and then by $\mathcal{L}_{|\lambda^*|} \subseteq \mathcal{L}_{\lceil |\lambda^*|/d \rceil} + \mathcal{L}_d$, which merely notes that a random variable can be always be decomposed into multiple constituent parts. Equivalently, any integer $[1, N]$ can be mapped injectively to an ordered tuple $\{[1, M], [1, \lceil N/M \rceil]\}$.

To be clear, it is an open question if Q_d is *ever* convex for finite d in scenarios where **Thm. (2)** does not apply. **Conj. 2** not only speculates the affirmative but effectively proposes a minimal value for what that finite d might be. Finally, note that an implication of **Conj. 1** and **Conj. 2** combined is that Q_d should be convex if and only if $d \geq |\lambda^*|$.

V. DISCUSSION

Questions remain even when still considering the $(2-2-2)$ scenario. We've established that $\mathcal{L}_\infty^{(2-2-2)} \not\subseteq Q_2$, but would

qutrits be able to span the local polytope, or would $d = 4$ be required?⁴ The insufficiency of qutrits is speculated in **Conj. 1**, but this should certainly be investigated further. Furthermore, although from **Cor. (3)** it is clear that $\mathcal{Q}_{16}^{(2-2-2)}$ is convex, there's still a large gap between the non-convex result of \mathcal{Q}_2 and the yes-convex result of \mathcal{Q}_{16} . The convexity of $\mathcal{Q}_4^{(2-2-2)}$ is speculated in **Conj. 2**, but this too should certainly be investigated further. For general scenarios where one presumes that **Thm. (2)** does not apply, such as (2-3-2) [78], we have noted it is a completely open question if \mathcal{Q}_d is *ever* convex for finite d .

Most generally, given three descriptive elements: 1) an operational description of some Bell scenario such as $(n-m-v)$, 2) the local Hilbert space dimension limit of d , and 3) the dimension limit of any shared classical randomness $|\lambda|$, are the resulting correlations $\mathcal{L}_{|\lambda|} + \mathcal{Q}_d$ convex? Such fundamental questions remain generally unanswered despite the broad results of **Thms. (1)** and **(2)**. The purely classical regime is considered further, however, in the Appendix B.

Quantifying the genuine boundaries of finite-dimensionally-generated quantum correlations, as opposed to the convex hull of such correlations, is an important question for future research. We have evidenced that nonconvexity should be expected in all nonlocality scenarios with sufficiently restrictive constraints on the local Hilbert space dimensions. As many physical systems have implicit bounds on their local Hilbert space dimensions it is all the more important to anticipate nonconvexity of quantum correlations in practical quantum information-theoretic protocols. Purely quantum systems with sufficiently low dimensionality are forbidden from displaying certain classical correlations, a property which may be exploited as a device-authenticating security check in quantum cryptographic implementations.

The nonconvexity of quantum correlations is no less relevant, practically, than other no-go results pertaining to finite Hilbert space dimensions. Quantifying the *genuine* boundaries of finite-dimensionally-generated quantum correlations, as opposed to the convex hull of such correlations [31, 32], is therefore an important question for future research. How the nonconvexity of dimensionally limited correlations may clarify the relationship between Hilbert space dimension and degree of classical randomness is an intriguing area for future work [83–85]. Perhaps classical information theory tools for considering limited shared randomness [86] can be adapted and applied to finite dimensional quantum systems. Fuller quantitative operational characterizations of finite dimension quantum systems are thus valuable desiderata for both foundational and practical research.

ACKNOWLEDGMENTS

We are grateful to Joshua Combes, Tobias Fritz, Ravi Kunjwal, Corsin Pfister, Marco Quintino, Sandu Popescu, Matt Pusey, Kevin Resch, Vincent Russo, and Rob Spekkens for valuable discussions, and to Matty Hoban for alerting us to a serious flaw in an early version of this manuscript. J. M. D. is grateful for support from the Natural Sciences and Engineering Research Council of Canada. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development and Innovation.

Appendix A: Correlations from asymmetric-dimensional quantum states

In Section II, we defined $\mathcal{Q}_d^{(n-m-v)}$ as the set of correlations achievable when all parties share a quantum state with *equal* local dimensions d . Here we note that any result concerning the uniform-dimension case also applies to the asymmetric-dimension case if the scenario in question is bipartite. We are grateful to referee comments for suggesting this extension.

Proposition (4): If, in some bipartite Bell scenario, Alice and Bob share a quantum system with local dimensions d_A and d_B , respectively, the set of distributions they can achieve is identical to the set of distributions achievable if the minimum dimension were common across both parties: $\mathcal{Q}_{\{d_A, d_B\}}^{(2-m-v)} = \mathcal{Q}_{d_{\min}}^{(2-m-v)}$, where $d_{\min} = \min\{d_A, d_B\}$.

Proof. In realizing achievable quantum distributions in a specified dimension, it suffices to consider pure states as the mixedness may be included in the local measurement settings; see [38, Lemma 1]. Any pure state in $\mathcal{H}^{d_A} \otimes \mathcal{H}^{d_B}$ is equivalent under local unitary transformations to its Schmidt decomposition, which has at most $d_{\min} = \min\{d_A, d_B\}$ terms [58–61]. Therefore, any correlation achievable with states in $\mathcal{H}^{d_A} \otimes \mathcal{H}^{d_B}$ may also be achieved with states in $\mathcal{H}^{d_{\min}} \otimes \mathcal{H}^{d_{\min}}$. \square

Note that this proof relies on the fact that the Schmidt decomposition of a bipartite state can be written as a sum of at most d_{\min} terms. This feature does not extend to general entanglement classes in multipartite scenarios [58]; in these scenarios, asymmetric-dimensional states may lead to unique sets of achievable distributions.

Appendix B: Bounding the Minimum Classical Shared Randomness which Spans the Local Polytope

In this section we establish some upper and lower bounds on $|\lambda^*|$, as defined per Eq. (4) in the main text. We are grateful to Matt Pusey of the Perimeter Institute for Theoretical Physics for suggesting most of these proofs. The related question of how to find an explicit

⁴ $|\lambda^*|^{(2-2-2)} = 4$ is proven in the Appendix B.

classical implementation of a box while minimizing $|\lambda|$ (“ontological compression”) is considered in Ref. [39].

Proposition (5) : $|\lambda^*| \leq v^{m(n-1)}$. In particular for the (2-2-2) scenario $|\lambda^*| \leq 4$.

Proof. Let $n-1$ parties have all their measurements depend deterministically on λ , and only the n ’th party depending probabilistically on λ . There are precisely $v^{m(n-1)}$ possible deterministic distributions among the $n-1$ parties, so without loss of generality it suffices to take $|\lambda| = v^{m(n-1)}$. In this manner every conceivable classical correlation is possible. \square

Proposition (6) : $|\lambda^*| \leq (m(v-1)+1)^n - 1$.

Proof. The proof is analogous to the proof of **Thm. (1)**. \mathcal{L}_1 is a pathwise-connected closed compactum of dimension $\mathcal{F}^{(n-m-v)} = (m(v-1)+1)^n - 1$ per Eq. (2). Fenchel’s theorem [72, 73] then implies that the Carathéodory number of any box \mathcal{L}_1 is less than or equal to \mathcal{F} ; i.e. any $P \in \text{ConvexHull}[\mathcal{L}_1]$ can be decomposed into a convex combination of at most \mathcal{F} boxes from \mathcal{L}_1 . $|\lambda^*|$ is identically that Carathéodory of \mathcal{L}_1 . \square

Proposition (7) : $|\lambda^*| \geq (m(v-1)+1)^{n-1} + \sum_{k=2}^n \sum_{j=2}^k \binom{n-1}{k-1} \binom{m}{j} S_2[k, j] (j-1)! (v-1)^k$, where S_2 refers to Stirling number of the second kind [52]. For $n=2$, therefore, $|\lambda^*| \geq \frac{m(m-1)(v-1)^2}{2} + m(v-1) + 1$ and in particular for the (2-2-2) scenario $|\lambda^*| \geq 4$.

Proof. Consider a box which is as random as possible while still satisfying

$$p[a+b+c\dots = 0 \bmod v \mid xyz] = 1 \quad \text{iff} \quad x = y = z\dots \quad (\text{B1})$$

i.e. the outputs of the parties always (modularly) sum to zero when the inputs are aligned, but this perfect correlation is not detected whenever the inputs are not all aligned. One can then, given $n-1$ of the outputs, perfectly determine the remaining one as $a_x = -b_x - c_x \dots \bmod v$. The instances of perfect correlation enforce that there is no local noise, and thus that every party’s output depends *deterministically* on the random variable. The degrees of freedom in the outputs of the first $n-1$ parties, given by Eq. (2) as $(m(v-1)+1)^{n-1} - 1$, are all able to be set independently and yet must be determined solely by the shared random variable. We must give the shared variable an alphabet size equal to the number of degrees of freedom *plus one*, to account for normalization. Thus, $|\lambda^*| \geq (m(v-1)+1)^{n-1}$.

This loose lower bound can be strengthened to the expression in **Prop. (7)** by counting the remaining degrees of freedom which include the last party’s outcome. There are $\binom{n-1}{k-1}$ ways to choose a k -partite context involving the last party. We then need to assign inputs to the parties; as the case where all inputs are equal is already specified

by the definition of the box, we only need to consider distributing $j > 1$ distinct inputs among the k parties. There are $\binom{m}{j}$ ways to choose which j distinct inputs to distribute. The Stirling number $S_2[k, j]$ gives the number of ways k objects can be divided into j (indistinguishable) partitions. Each partition is assigned one input however, so we must consider the various permutations of mapping inputs to partitions. To avoid over-specifying the distribution we take the last party (and any other parties in the partition including the last party) to be associated with largest of the j selected inputs. Thus we only consider $(j-1)!$ permutations of how the smaller inputs can be assigned to the remaining partitions. Finally we specify the outputs, avoiding output 0 per the parameterization scheme discussed prior to Eq. (2). Thus there are $v-1$ distinct outputs considered for each of the k parties considered. \square

For pedagogical clarity we demonstrate how to obtain $|\lambda^*| \geq 7$ in this fashion for the (2-3-2) scenario, famous for the Bell inequality I_{3322} [78–80]. The classical box in (2-3-2) which satisfies $b_x = a_x$ has six degrees of freedom, namely $p[a_0=1]$, $p[a_1=1]$, $p[a_2=1]$, $p[a_1=1, b_1=1]$, $p[a_0=1, b_2=1]$, and $p[a_1=1, b_2=1]$. Trivially $p[a_x=0] = 1 - p[a_x=1]$ and $p[a_x=1, b_y=0] = p[a_x=1] - p[a_x=1, b_y=1]$ etc. Note that $p[b_y=i]$ is fixed as equal to $p[a_y=i]$. Furthermore, $p[a_x=i, b_y=j] = p[a_y=j, b_x=i]$, so we only ever enumerate $p[a_x=i, b_y=j]$ where $x < y$. Adding one to the six degrees of freedom yields $|\lambda^*| \geq 7$.

Appendix C: A Complimentary Fundamental Axiom

In **Ax. (1)**, it was noted that the set of boxes achievable with classical shared randomness of dimension $|\lambda|$ is achievable with quantum states of dimension $d = |\lambda|$, and furthermore, realizable with separable quantum states. Here, we briefly note that while the inverse is not true, it *is* true that any box achievable with separable quantum states of dimension d can be achieved with classical shared randomness of finite dimension.

Axiom (3) : $(\mathcal{Q} : \text{sep})_d \subseteq \mathcal{L}_{|\lambda| \geq d^n}$, i.e. all quantum correlations generated by separable states are also classically achievable given enough shared randomness, where enough means $|\lambda| \geq d^n$.

Corollary (4) : Entanglement is required for non-locality.

Proof. Recall that by definition separable states can be written as $\rho_{\text{sep}} = \sum_i c^{(i)} \rho_A^{(i)} \otimes \rho_B^{(i)} \dots$, and therefore satisfy $\text{Tr}[\rho_{\text{sep}} (\hat{A}_{a|x} \otimes \hat{B}_{b|y} \dots)] = \sum_{i=1}^{i_{\text{max}}} (c^{(i)} \text{Tr}[\rho_A^{(i)} \hat{A}_{a|x}] \text{Tr}[\rho_B^{(i)} \hat{B}_{b|y}] \dots)$. Any separable state with local Hilbert space dimension d can be decomposed into a mixture of no-more-than d^n product states [53 Def. 6, 54 Thm. 2] so we can replace i_{max} with d^n without loss of generality. One can therefore

map every $P \in (\mathcal{Q} : \text{sep})_d$ to a $P' \in \mathcal{L}_{|\lambda|=d^n}$ by the following construction on Eq. (3): $\lambda \rightarrow i$, $p[\lambda] \rightarrow c^{(i)}$, $p[a|x\lambda] \rightarrow \text{Tr}[\rho_A^{(i)} \hat{A}_{a|x}]$, etc. \square

Goh *et al.* [67] have shown that merely $|\lambda| \geq d$ is not always sufficient to classically simulate the correlations which result from separable states, i.e. that separable states can still manifest super-locality. In particular, they show that the box $p[ab|xy] = \frac{2+(-1)^{a+b+xy}\sqrt{2}}{8}$ requires $|\lambda| > 2$, but can nevertheless be achieved by measurements on separable qubits.

Appendix D: Details of the Parameterization of Shared-Qubits Boxes in the (2-2-2) Scenario

We consider qubit-based boxes with explicit representations in terms of two-qubit states of arbitrary entanglement and general two-outcome POVMs. We need to consider exclusively pure states, per Ref. [38, Lemma 1]. As all pure states are equivalent to their Schmidt-decomposed form under local unitary transformations [58–61], it is thus sufficient to consider the state

$$|\psi\rangle = \cos(\frac{\alpha}{2})|00\rangle + \sin(\frac{\alpha}{2})|11\rangle, \quad \alpha \in (0, \pi), \quad (\text{D1})$$

by folding the local degrees of freedom into the measurement operators.

While POVMs can be regarded as projective measurements in a larger Hilbert space, restricting the dimensionality of the quantum systems necessitates the use of general POVMs [41, 55]. Adapting the notation of [62], we express a general binary 0/1 outcome POVM element as

$$\hat{A}_{a|x} = \frac{1}{2} [(1 + (-1)^{1-a}\kappa_{A_x}) \mathbb{1} + (-1)^{1-a}\eta_{A_x} (\vec{n}_{A_x} \cdot \vec{\sigma})], \quad (\text{D2})$$

where $\vec{\sigma} = (\hat{\sigma}_X, \hat{\sigma}_Y, \hat{\sigma}_Z)$ is a vector of Pauli matrices and $\vec{n}_{A_x} = (\sin \theta_{A_x} \cos \phi_{A_x}, \sin \theta_{A_x} \sin \phi_{A_x}, \cos \theta_{A_x})$ is a unit vector defining a direction in the Bloch sphere in spherical coordinates. Bob's POVM elements are defined similarly in a separate Hilbert space. Technically, the measurement operators of the distinct parties need only commute with each other to form a genuine quantum multipartite implementation; we relegate each party to a distinct Hilbert space for convenience to ensure appropriate commutativity, which has apparently no loss of generality [17, 63–65]. To ensure positivity of the POVM elements corresponding to both outputs, the following conditions must be met:

$$\forall_{A_x} : \eta_{A_x} - 1 \leq \kappa_{A_x} \leq 1 - \eta_{A_x}. \quad (\text{D3})$$

Note that Eq. (D3) implies $0 \leq \eta_{k_x} \leq 1$. If both bounds in Eq. (D3) are simultaneously saturated then $\kappa_{k_x} = 0$ and $\eta_{k_x} = 1$, and Eq. (D2) represents a projection-valued measurement (PVM).

As we are concerned with two-outcome POVMs, it is conventional to parameterize boxes in terms of bias of the measurement outcome, i.e.

$$\langle A_x \rangle = \langle \hat{A}_{1|x} \rangle - \langle \hat{A}_{0|x} \rangle \quad (\text{D4})$$

$$\begin{aligned} \langle A_x B_y \rangle &= \langle \hat{A}_{1|x} \hat{B}_{1|y} \rangle - \langle \hat{A}_{0|x} \hat{B}_{1|y} \rangle \\ &\quad - \langle \hat{A}_{1|x} \hat{B}_{0|y} \rangle + \langle \hat{A}_{0|x} \hat{B}_{0|y} \rangle. \end{aligned} \quad (\text{D5})$$

The four marginal and four joint biases (for all x and y options) parameterize the eight-dimensional conditional probability space for the (2-2-2) scenario, and relate to the 17 “backend” parameters of the state and measurements as

$$\langle A_x \rangle = \eta_{A_x} \cos(\alpha) \cos(\theta_{A_x}) + \kappa_{A_x}, \text{ and} \quad (\text{D6})$$

$$\begin{aligned} \langle A_x B_y \rangle &= \eta_{A_x} \eta_{B_y} \cos(\phi_{A_x} + \phi_{B_y}) \sin(\alpha) \sin(\theta_{A_x}) \sin(\theta_{B_y}) \\ &\quad + \eta_{A_x} \eta_{B_y} \cos(\theta_{A_x}) \cos(\theta_{B_y}) \\ &\quad + \eta_{A_x} \kappa_{B_y} \cos(\alpha) \cos(\theta_{A_x}) \\ &\quad + \eta_{B_y} \kappa_{A_x} \cos(\alpha) \cos(\theta_{B_y}) + \kappa_{A_x} \kappa_{B_y}. \end{aligned} \quad (\text{D7})$$

Appendix E: Convexly Combining Quantum Boxes via Direct Sum of Hilbert Spaces

Suppose we wish to take the convex combination of N qudit-based multipartite boxes, i.e.

$$\tilde{P} = \sum_{i=0}^{N-1} c_i P_i \quad \text{where} \quad P_i \in \mathcal{Q}_d. \quad (\text{E1})$$

Typically⁵, $\tilde{P} \notin \mathcal{Q}_d$. This can be built as follows:

Index the local measurement operators of each of the N boxes being combined into \tilde{P} by $\hat{A}_{a|x}^{(i)}, \hat{B}_{b|y}^{(i)}$ etc. Now imagine – although entirely unjustified – that all the N boxes P_i are predicated on sharing the *same* composite quantum state, i.e. $\forall_i \rho_i = \rho_0$. In order to reproduce the marginal probabilities of \tilde{P} with a single quantum box (without supplementary shared randomness), we should take $\tilde{\rho} = \rho_0$ and define the new measurement operators $\hat{A}_{a|x} = \sum_{i=0}^{N-1} c_i \hat{A}_{a|x}^{(i)}$. This satisfies the requirement that $\tilde{p}[a|x] = \sum_{i=0}^{N-1} c_i p_i[a|x]$ etc.

Unfortunately, choosing measurement operators to satisfy the single-partite marginal probabilities does not extend to satisfaction of the required bipartite joint probabilities. Following Eq. (1) we find that

$$\begin{aligned} p[ab|xy] &= \text{Tr}[\tilde{\rho} \hat{A}_{a|x} \otimes \hat{B}_{b|y}] \\ &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} c_i c_j \text{Tr}[\rho_0 \hat{A}_{a|x}^{(i)} \otimes \hat{B}_{b|y}^{(j)}] \end{aligned} \quad (\text{E2})$$

⁵ The major exception is when \mathcal{Q}_d is convex, as then $\tilde{P} \in \mathcal{Q}_d$ by definition. Convexity might be “rare”, however.

which does not remotely match $\tilde{p}[ab|xy] = \sum_i c_i p_i[ab|xy] = \sum_i c_i \text{Tr}[\rho_0 \hat{A}_{a|x}^{(i)} \otimes \hat{B}_{b|y}^{(i)}]$. Therefore quantum boxes predicated on different sets of local measurement operators cannot be combined without increasing the Hilbert space dimension⁶; as per **Ax. (2)**, however, $\tilde{P} \in \mathcal{Q}_{d \times N}$.

Imagine that each P_i is implemented in $\mathcal{H}^{(i)} = (\mathbb{C}^d)^{\otimes n}$ using the (distinct!) quantum states $\rho^{(i)}$ and local measurement operators $\hat{A}_x^{(i)}$, $\hat{B}_y^{(i)}$, etc. Then, to implement \tilde{P} using a single quantum box, one may take the direct sum of the component Hilbert spaces such that

$$\tilde{\mathcal{H}} = \bigoplus_{i=0}^{N-1} \mathcal{H}^{(i)}, \quad \tilde{\rho} = \bigoplus_{i=0}^{N-1} c_i \rho_i, \quad \widetilde{\hat{A}_{a|x}} = \bigoplus_{i=0}^{N-1} \hat{A}_{a|x}^{(i)}, \quad (E3)$$

and so on.

To explain how Eq. (E3) automatically satisfies Eq. (E1), and why the new local Hilbert space dimension in Eq. (E3) can be thought of as $d \times N$, consider how \tilde{P} is implemented in $((\mathbb{C}^d)^{\otimes N})^{\otimes n}$. The idea is to assign an ancilla \mathbb{C}^N to every party, and to make the new shared state simultaneously diagonalized in all the ancillae. Thus,

$$\begin{aligned} \tilde{\rho} &= \sum_{i=0}^{N-1} c_i \rho_i \otimes |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B \otimes \dots, \\ \widetilde{\hat{A}_{a|x}} &= \sum_{i=0}^{N-1} \hat{A}_{a|x}^{(i)} \otimes |i\rangle\langle i|_A, \\ \widetilde{\hat{B}_{b|y}} &= \sum_{j=0}^{N-1} \hat{B}_{b|y}^{(j)} \otimes |j\rangle\langle j|_B, \end{aligned} \quad (E4)$$

and so on. Indeed, Eq. (E4) amounts to the *definition* of the direct sum in Eq. (E3). Thus the parties' new measurement operators now live in distinct Hilbert spaces N times larger than originally, and the new shared state is a cq-state [40, 44], namely a block-diagonal composition of the component original states in the convex combination per Eq. (E1).

What we find is that *direct summation of Hilbert spaces is identically convexification* in the sense of Eq. (5). Since $\bigoplus^N \bigotimes^n \mathbb{C}^d \subset \bigotimes^N \bigotimes^n \mathbb{C}^d$, we have proven that $\mathcal{L}_N + \mathcal{Q}_d \subseteq \mathcal{Q}_{d \times N}$.

⁶ Suppose we artificially consider global measurements of the form $\hat{A}''\hat{B}'' = \frac{\hat{A} \otimes \hat{B} + \hat{A}' \otimes \hat{B}'}{2}$. Such not-local-but-still-separable measurements lack physical meaning, but they can be interpreted as representing the convex combination of boxes acting on the same shared state with different local, i.e. product state, measurements. Such artificial global measurements are known as *separable superoperators*, and have been studied elsewhere in the context of state discrimination [68, 69].

-
- [1] J. S. Bell, “On the Einstein-Podolsky-Rosen paradox,” *Physics* **1**, 195 (1964).
 - [2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, “Bell nonlocality,” *Rev. Mod. Phys.* **86**, 419 (2014).
 - [3] S. Popescu, “Nonlocality beyond quantum mechanics,” *Nat. Phys.* **10**, 264 (2014).
 - [4] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, “Nonlocal correlations as an information-theoretic resource,” *Phys. Rev. A* **71**, 022101 (2005).
 - [5] J. Barrett and S. Pironio, “Popescu-Rohrlich correlations as a Unit of nonlocality,” *Phys. Rev. Lett.* **95**, 140401 (2005).
 - [6] N. S. Jones and L. Masanes, “Interconversion of nonlocal correlations,” *Phys. Rev. A* **72**, 052312 (2005).
 - [7] J.-D. Bancal, C. Branciard, N. Brunner, N. Gisin, and Y.-C. Liang, “A framework for the study of symmetric full-correlation Bell-like inequalities,” *J. Phys. A* **45**, 125301 (2012).
 - [8] V. Scarani, “Quantum information: Primitive notions and quantum correlations,” [arXiv:0910.4222](https://arxiv.org/abs/0910.4222) (2009).
 - [9] V. Scarani, “The device-independent outlook on quantum physics,” *Acta Physica Slovaca* **62**, 347 (2012).
 - [10] J.-D. Bancal, *On the Device-Independent Approach to Quantum Physics* (Springer International Publishing, 2014).
 - [11] G. Chiribella and X. Yuan, “Bridging the gap between general probabilistic theories and the device-independent framework for nonlocality and contextuality,” [arXiv:1504.02395](https://arxiv.org/abs/1504.02395) (2015).
 - [12] A. Acín, S. Massar, and S. Pironio, “Efficient quantum key distribution secure against no-signalling eavesdroppers,” *New J. Phys.* **8**, 126 (2006).
 - [13] L. Masanes, S. Pironio, and A. Acín, “Secure device-independent quantum key distribution with causally independent measurement devices,” *Nat. Comm.* **2**, 238+ (2011).
 - [14] A. Ekert and R. Renner, “The ultimate physical limits of privacy,” *Nature* **507**, 443 (2014).
 - [15] R. F. Werner and M. M. Wolf, “All-multipartite bell-correlation inequalities for two dichotomic observables per site,” *Phys. Rev. A* **64**, 032112 (2001).
 - [16] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, “A combinatorial approach to nonlocality and contextuality,” *Comm. Math. Phys.* **334**, 533 (2015).
 - [17] T. Fritz, “Tsirelson’s problem and Kirchberg’s conjecture,” *Rev. Math. Phys.* **24**, 1250012 (2012).
 - [18] N. Ozawa, “Tsirelson’s problem and asymptotically commuting unitary matrices,” *J. Math. Phys.* **54**, 032202 (2013).
 - [19] D. Avis and T. Ito, “Comparison of two bounds of the quantum correlation set,” in *1st Inter. Conf. on Quant. Nano & Micro Tech.* (IEEE, 2007).
 - [20] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, “Almost quantum correlations,” *Nat. Commun.* **6**, 6288 (2015).
 - [21] M. Navascués, S. Pironio, and A. Acín, “Bounding the set of quantum correlations,” *Phys. Rev. Lett.* **98**, 010401 (2007).
 - [22] M. Navascués, S. Pironio, and A. Acín, “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations,” *New J. Phys.* **10**, 073013 (2008).
 - [23] S. Pironio, M. Navascués, and A. Acín, “Convergent relaxations of polynomial optimization problems with non-commuting variables,” *SIAM J. Optim.* **20**, 2157 (2010).
 - [24] D. Pérez-García, M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge, “Unbounded violation of tripartite Bell inequalities,” *Comm. Math. Phys.* **279**, 455 (2008).
 - [25] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf, “Operator space theory: A natural framework for Bell inequalities,” *Phys. Rev. Lett.* **104**, 170405 (2010).
 - [26] M. Epping, H. Kampermann, and D. Bruß, “Designing Bell inequalities from a Tsirelson bound,” *Phys. Rev. Lett.* **111**, 240404 (2013).
 - [27] M. Żukowski and Č. Brukner, “Bell’s theorem for general N -qubit states,” *Phys. Rev. Lett.* **88**, 210401 (2002).
 - [28] I. Devetak and A. Winter, “Distillation of secret key and entanglement from quantum states,” *Proc. Roy. Soc. A* **461**, 207 (2005).
 - [29] B. Kraus, N. Gisin, and R. Renner, “Lower and upper bounds on the secret-key rate for quantum key distribution protocols using one-way classical communication,” *Phys. Rev. Lett.* **95**, 080501 (2005).
 - [30] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. A. Méthot, and V. Scarani, “Testing the dimension of Hilbert spaces,” *Phys. Rev. Lett.* **100**, 210503 (2008).
 - [31] M. Navascués, G. de la Torre, and T. Vértesi, “Characterization of quantum correlations with local dimension constraints and its device-independent applications,” *Phys. Rev. X* **4**, 011011 (2014).
 - [32] M. Navascués and T. Vértesi, “Bounding the set of finite dimensional quantum correlations,” *Phys. Rev. Lett.* **115**, 020501 (2015).
 - [33] J. Bowles, M. T. Quintino, and N. Brunner, “Certifying the dimension of classical and quantum systems in a prepare-and-measure scenario with independent devices,” *Phys. Rev. Lett.* **112**, 140407 (2014).
 - [34] R. Gallego, N. Brunner, C. Hadley, and A. Acín, “Device-independent tests of classical and quantum dimensions,” *Phys. Rev. Lett.* **105**, 230501 (2010).
 - [35] R. Jain, Y. Shi, Z. Wei, and S. Zhang, “Efficient protocols for generating bipartite classical distributions and quantum states,” *IEEE Trans. Info. Theo.* **59**, 5171 (2013).
 - [36] H. Fawzi, J. Gouveia, P. A. Parrilo, R. Z. Robinson, and R. R. Thomas, “Positive semidefinite rank,” *Math. Program.* **153**, 133 (2015).
 - [37] S. Zhang, “Quantum strategic game theory,” in *Proc. 3rd Innov. Theo. Comput. Sci. - ITCS ’12* (ACM, New York, NY, USA, 2012) pp. 39–59, 429122.
 - [38] J. Sikora, A. Varvitsiotis, and Z. Wei, “On the minimum dimension of a Hilbert space needed to generate a quantum correlation,” [arXiv:1507.00213](https://arxiv.org/abs/1507.00213) (2015).
 - [39] N. Harrigan, T. Rudolph, and S. Aaronson, “Representing probabilistic data via ontological models,” [arXiv:0709.1149](https://arxiv.org/abs/0709.1149) (2007).
 - [40] S. Wehner, M. Christandl, and A. C. Doherty, “Lower bound on the dimension of a quantum system given measured data,” *Phys. Rev. A* **78**, 062112 (2008).

- [41] K. F. Pál and T. Vértesi, “Concavity of the set of quantum probabilities for any given dimension,” *Phys. Rev. A* **80**, 042114 (2009).
- [42] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden-variable theories,” *Phys. Rev. Lett.* **23**, 880 (1969).
- [43] E. Wolfe and S. F. Yelin, “Quantum bounds for inequalities involving marginal expectation values,” *Phys. Rev. A* **86**, 012123 (2012).
- [44] M. Piani, P. Horodecki, and R. Horodecki, “No-local-broadcasting theorem for multipartite quantum correlations,” *Phys. Rev. Lett.* **100**, 090502 (2008).
- [45] A. Acín, R. Gill, and N. Gisin, “Optimal Bell tests do not require maximally entangled states,” *Phys. Rev. Lett.* **95**, 210402 (2005).
- [46] S. Pironio, “Lifting Bell inequalities,” *J. Math. Phys.* **46**, 062112 (2005).
- [47] A. Fine, “Hidden variables, joint probability, and the Bell inequalities,” *Phys. Rev. Lett.* **48**, 291 (1982).
- [48] R. Spekkens, “The status of determinism in proofs of the impossibility of a noncontextual model of quantum theory,” *Found. Phys.* **44**, 1125 (2014).
- [49] R. Kunjwal, “Fine’s theorem, noncontextuality, and correlations in Specker’s scenario,” *Phys. Rev. A* **91**, 022108 (2015).
- [50] I. Bárány, “A generalization of Carathéodory’s theorem,” *Discrete Mathematics* **40**, 141 (1982).
- [51] E. W. Weisstein, “Carathéodory’s fundamental theorem,” (2015), see also mathoverflow.net/q/77379.
- [52] E. W. Weisstein, “Stirling number of the second kind,” (2015), see also math.stackexchange.com/a/264814.
- [53] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of n -particle mixed states: necessary and sufficient conditions in terms of linear maps,” *Phys. Lett. A* **283**, 1 (2001).
- [54] R. Lockhart, “Optimal ensemble length of mixed separable states,” *J. of Math. Phys.* **41**, 6766 (2000).
- [55] T. Vértesi and E. Bene, “Two-qubit Bell inequality for which positive operator-valued measurements are relevant,” *Phys. Rev. A* **82**, 062115 (2010).
- [56] B. Groisman, S. Popescu, and A. Winter, “Quantum, classical, and total amount of correlations in a quantum state,” *Phys. Rev. A* **72**, 032317 (2005).
- [57] B. S. Cirel’son, “Quantum generalizations of Bell’s inequality,” *Lett. Math. Phys.* **4**, 93 (1980).
- [58] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach, “Generalized Schmidt decomposition and classification of three-quantum-bit states,” *Phys. Rev. Lett.* **85**, 1560 (2000).
- [59] D. Bruß, “Characterizing entanglement,” *J. Math. Phys.* **43**, 4237 (2002).
- [60] L. Amico, A. Osterloh, and V. Vedral, “Entanglement in Many-Body Systems,” *Rev. Mod. Phys.* **80**, 517 (2008).
- [61] O. Gühne and G. Tóth, “Entanglement detection,” *Phys. Rep.* **474**, 1 (2009).
- [62] T. Heinosaari, D. Reitzner, and P. Stano, “Notes on joint measurability of quantum observables,” *Found. Phys.* **38**, 1133 (2008).
- [63] M. Junge, M. Navascués, C. Palazuelos, D. Perez-Garcia, V. B. Scholz, and R. F. Werner, “Connes’s embedding problem and Tsirelson’s problem,” *J. Math. Phys.* **52**, 012102 (2011).
- [64] M. Navascués, T. Cooney, D. Pérez-García, and N. Villanueva, “A physical approach to Tsirelson’s problem,” *Found. Phys.* **42**, 985 (2012).
- [65] V. B. Scholz and R. F. Werner, “Tsirelson’s problem,” [arXiv:0812.4305](https://arxiv.org/abs/0812.4305) (2008).
- [66] See Supplemental Material in arXiv source at [arXiv:1506.01119](https://arxiv.org/abs/1506.01119).
- [67] K. T. Goh, J.-D. Bancal, and V. Scarani, ““certification of entanglement based only on hilbert space dimension”,” [arXiv:1509.08682](https://arxiv.org/abs/1509.08682) (2015).
- [68] S. Bandyopadhyay and M. Nathanson, “Tight bounds on the distinguishability of quantum states under separable measurements,” *Phys. Rev. A* **88**, 052313 (2013).
- [69] S. Bandyopadhyay, A. Cosentino, N. Johnston, V. Russo, J. Watrous, and N. Yu, “Limitations on separable measurements by convex optimization,” [arXiv:1408.6981](https://arxiv.org/abs/1408.6981) (2014).
- [70] J. Baez, “Tensor product and direct sum,” See also [ProofWiki:Hilbert.Space.Direct.Sum](https://proofwiki.org/wiki/Hilbert_Space_Direct_Sum).
- [71] J. K. Hunter and B. Nachtergaele, *Applied Analysis* (World Scientific, 2001).
- [72] O. Hanner and H. Rådström, “A generalization of a theorem of Fenchel,” *Proc. Amer. Math. Soc.* **2**, 589 (1951).
- [73] I. Bárány and R. Karasev, “Notes about the Carathéodory number,” *Disc. Comp. Geom.* **48**, 783 (2012).
- [74] L. Masanes, “Extremal quantum correlations for N parties with two dichotomic observables per site,” [quant-ph/0512100](https://arxiv.org/abs/quant-ph/0512100) (2005).
- [75] L. Masanes, “Asymptotic violation of Bell inequalities and distillability,” *Phys. Rev. Lett.* **97**, 050503 (2006).
- [76] T. Fritz, “Polyhedral duality in Bell scenarios with two binary observables,” *J. Math. Phys.* **53**, 072202 (2012), [10.1063/1.4734586](https://doi.org/10.1063/1.4734586).
- [77] D. Collins and N. Gisin, “A Relevant Two Qubit Bell Inequality Inequivalent to the CHSH Inequality,” *J. Phys. A* **37**, 1775 (2004).
- [78] K. F. Pál and T. Vértesi, “Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems,” *Phys. Rev. A* **82**, 022116 (2010).
- [79] Y.-C. Liang, T. Vértesi, and N. Brunner, “Semi-device-independent bounds on entanglement,” *Phys. Rev. A* **83**, 022108 (2011).
- [80] C. Palazuelos and Z. Yin, “Large bipartite Bell violations with dichotomic measurements,” [arXiv:1504.05769](https://arxiv.org/abs/1504.05769) (2015).
- [81] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, “Bell inequalities for arbitrarily high-dimensional systems,” *Phys. Rev. Lett.* **88**, 040404 (2002).
- [82] S. Zohren and R. D. Gill, “Maximal violation of the Collins-Gisin-Linden-Massar-Popescu inequality for infinite dimensional states,” *Phys. Rev. Lett.* **100**, 120406 (2008).
- [83] A. Acín, S. Massar, and S. Pironio, “Randomness versus nonlocality and entanglement,” *Phys. Rev. Lett.* **108**, 100402 (2012).
- [84] A. A. Methot and V. Scarani, “An anomaly of non-locality,” *Quant. Info. Comp.* **7**, 157 (2007).
- [85] A. Plastino, G. Bellomo, and A. R. Plastino, “Quantum state space-dimension as a quantum resource,” [arXiv:1505.05455](https://arxiv.org/abs/1505.05455) (2015).
- [86] R. Chaves, L. Luft, and D. Gross, “Causal structures from entropic information: geometry and novel scenarios,” *New J. Phys.* **16**, 043001 (2014).